

1/22

Figure 1

1 GAATTCGCGGCCGCGTCGACAGATGCCTTCTTCTGCCTGAGATTACACCCCACTAGCCAA 60

61 CCACTTTTGCCTTCGAATGGAGAGATCCAGGTACGGGAAGAACCGGGCAGCTCACCTGG 120

121 ACCCGACTGCCCAAGGGTTCAAGAACTCCCCGACCATCTTGACGAAGCCCTACACAGG 180

181 GACCTGGCCAACCTCAGGATCCAACACCCCTCAGGTGACCCCTCCTCCAGTACGTGGATGAC 240

241 CTGCTTCTGGCGGGAGCCACCAAACAGGACTGCTTAGAAGGTACGAAGGCACTAGCTG 300

301 GAATTGTCTGACCTAGGCTACAGAGCCTCTGCTAAGAAGGCCAGATTGCAGGAGAGAG 360

361 GTAACATACTGGGTACAGTTGCGGGGCGGGCAGCGATGGCTGACGGAGGCACGGAAG 420

421 AAAACTGTAGTCCAGATACCGGCCCCAACCACAGCCAAACAAGTGAGAGAGTTTGGGG 480

481 ACAGCTGGATTTGCAGACTGTGGATCCCGGGTTGCGACCTTAGCAGCCCCACTCTAC 540

541 CCGCTAACCAAAGAAAAAGGGGGATTCTCCTGGCCTGAGCACCAGAAGGCATTGAT 600

601 GCTATCAAAAGGCCCTGCTGAGGGCACCTGCTCTGGCCCTCCCTGACGTAACAAACCC 660

661 TTTACCCCTTATGTGGATGAGCGTAAGGGAGTAGCCCGAGGAGTTAACCCAAACCTA 720

721 GGACCATGGAGGAGACCTGTTGCCTACCTGTCAAAGAAGCTTGATCCTGTAGCCAGTGGT 780

781 TGGCCCGTATGTCTGAAGGCTATCGCAGCTGTGGCCATACTGGTCAAGGACGCTGACAAA 840

841 TTGACTTTGGACAGAATATAACTGTAATAGCCCCCATGCATTGGAGAACATCGTTGG 900

901 CAGCCCCCAGACCGATGGATGACCAACGCCCGCATGACCCACTATCAAAGCCTGCTTCTC 960

961 ACAGAGAGGGTCACTTGCTCCACCAGCCGCTCTCAACCCCTGCCACTCTTCTGCCTGAA 1020

1021 GAGACTGATGAACCAGTGACTCATGATTGCCATCAACTATTGATTGAGGAGACTGGGTC 1080

1081 CGCAAGGACCTTACAGACATAACCGCTGACTGGAGAAGTGCTAACCTGGTTCACTGACGGA 1140

1141 AGCAGCTATGTGGTGGAAAGGTAAAGAGGATGGCTGGGGCGGCAGTGGTGGACGGGACCCGC 1200

1201 ACGATCTGGGCCAGCAGCCTGCCGGAGGAACCTCAGCGAAAAGGCTGAGCTCATGGCC 1260

Figure 1 cont.

1261	CTCACGCAAGCTTGC GGCTGGCGAAGGGAAATCCATAAACATTTACGGACAGCAGG	1320
1321	TATGCCTTGCGACTGCACACGTACACGGGGCATCTATAAACAAAGGGGTTGCTTACC	1380
1381	TCAGCAGGGAGGGAAATAAAGAACAAAGAGGAAATTCTAAGCCTATTAGAAGCCTACAT	1440
1441	TTGCCAAAAAGGCTAGCTATTATACACTGTCCTGGACATCAGAAAGCCAAAGATCTCATA	1500
1501	TCTAGAGGGAACAGATGGCTGACCGGGTTGCCAAGCAGGCAGGCCAGGCTGTTAACCTT	1560
1561	CTGCCTATAATAGAAACGCCAAAGCCCAGAACCCAGACGACAGTACACCCCTAGAACAGAC	1620
1621	TGGCAAGAGATAAAAAAGATAGACCAGTTCTCTGAGACTCCGGAGGGACCTGCTATACC	1680
1681	TCATATGGGAAGGAAATCCTGCCCAACAAAGAAGGGTTAGAATATGTCCAACAGATACT	1740
1741	CGTCTAACCCACCTAGGAAC TAAACACCTGCAGCAGTTGGTCAGAACATCCCCTTATCAT	1800
1801	GTTCTGAGGCTACCAGGGAGTGGCTGACTCGGTGGTCAAACATTGTGTGCCCTGCCAGCTG	1860
1861	GTAAATGCTAATCCTCCAGAATACCTCCAGGAAAGAGACTAAGGGGAAGCCACCCAGGC	1920
1921	GCTCACTGGGAAGTGGACTTCACTGAGGTAAAGCCGGCTAAATACGGAAACAAATATCTA	1980
1981	TTGGTTTTGTAGACACCTTTCAGGATGGTAGAGGGCTTATCCTACTAAGAAAGAGACT	2040
2041	TCAACCGTGGTGGCTAAGAAAATCTGGAGGAAATTTCAGGAAAGATTTGGAAATACCTAAG	2100
2101	GTAATAGGGTCAGACAATGGTCCAGCTTCGTTGCCAGGTAAGTCAGGGACTGCCAAG	2160
2161	ATATTGGGGATTGATTGGAAACTGCATTGTGCATAAGACCCAAAGCTCAGGACAGGTA	2220
2221	GAGAGGATGAATAGAACCATTAAGAGACCCCTACCAAATTGACCACAGAGACTGGCATT	2280
2281	AATGATTGGATGGCTCTCCTGCCCTTGTGCTTTAGGGTAGGAACACCCCTGGACAG	2340
2341	TTTGGGCTGACCCCTATGAATTGCTCTACGGGGACCCCCCGTTGGCAGAAATTGCC	2400
2401	TTTGCACATAGTGCTGATGTGCTGCTTCCAGCCTTGTCTAGGCTCAAGGCGCTC	2460
2461	GAGTGGGTGAGGCAGCGAGCGTGGAAAGCAGCTCCGGAGGCCTACTCAGGAGGAGACTTG	2520

Figure 1 cont.

2521	CAAGTTCCACATCGCTTCCAAGTTGGAGATTAGTCTATGTTAGACGCCACCGTCAGGA	2580
2581	AACCTCGAGACTCGGTGGAAGGGACCTTATCTGTACTTTGACCACACCAACGGCTGTG	2640
2641	AAAGTCGAAGGAATCCCCACCTGGATCCATGCATCCCACGTTAAGCCGGCGCCACCTCCC	2700
2701	GATTGGGGTGGAAAGCCGAAAAGACTGAAAATCCCTTAAGCTTCGCCTCCATCGCTG	2760
2761	GTTCCCTTACTCTGTCAATAACTCCTCAAGTTAATGGTAAACGCCCTGTGGACAGCCCGAA	2820
2821	CTCCCATAAACCTTATCTCACCTGGTTACTTACTGACTCCGGTACAGGTATTAAATAT	2880
2881	TAACAGCACTCAAGGGGAGGCTCCCTGGGGACCTGGTGGCCTGAATTATATGTCTGCCT	2940
2941	TCGATCAGTAATCCCTGGTCTCAATGACCAGGCCACACCCCCCGATGTACTCCGTGCTTA	3000
3001	CGGGTTTACGTTGCCAGGACCCCCAAATAATGAAGAATATTGTGGAAATCCTCAGGA	3060
3061	TTTCCTTGCAAGCAATGGAGCTGCATAACTTCTAATGATGGGAATTGGAAATGCCAGT	3120
3121	CTCTCAGCAAGACAGAGTAAGTTACTCTTTGTTAACAAATCCTACCAAGTTATAATCAATT	3180
3181	TAATTATGCCATGGAGATGGAAAGATTGGCACAGCGGGTACAAAAAGATGTACGAAA	3240
3241	TAAGCAAATAAGCTGTCATTGTTAGACCTAGATTACTTAAAAATAAGTTCACTAAAAAA	3300
3301	AAAAAAAAAAAAAAAAAAAAA 3320	

4/22

Figure 2

1 TGTGGCCCCAGCGCGCTTGGAAATAAAATCCTCTTGCTGTTGCATCAAGACCGCTTCT 60

61 CGTGAGTGATTTGGGGTGTGCCTCTTCCGAGCCCGGACGAGGGGGATTGTTCTTTACT 120

121 GGCCTTCATTGGTGCCTGGCCGGAAATCCTGCGACCACCCCTAACACCGAGAAC 180

181 GACTTGGAGGTAAAGGGATCCCCTTGGAACGTGTGTGTGTCGGCCGGCGTCTGT 240

241 CTGAGTGTCTGTTTGGTGAATGCGCGCTTCGGTTGCAGCTGTCTCTCAGACCGTAA 300

301 GGACTGGAGGACTGTGATCAGCAGACGTGCTAGGAGGATCACAGGCTGCCACCCGGGG 360

361 ACGCCCCGGAGGTGGGAGAGGCCAGGGACGCCCTGGTGGTCTCCTACTGTGGTCAGAGG 420

421 ACCGAGTTCTGTTGTAAGCGAAAGCTTCCCCCTCCGGCCGTCCGACTCTTTGCCT 480

481 GCTTGTGGAAGACCGGGACGGGTCGCGTGTGTCTGGATCTGTTGGTTCTGTCTCGTGT 540

541 TCTTTGTCTTGTGCCTTGCTACAGTTAATATGGGACAGACAGTGAATACCCCCC 600

601 TTAGTTGACTCTGACCATTGGACTGAAGTTAGATCCAGGGCTCATATTGTCAAGTT 660

661 AGGTTAAGAAGGGACCTTGGCAGACTTCTGTGCCTCTGAATGGCAAACATTGATGTTG 720

721 GATGGCCATCAGAGGGGACCTTAATTCTGAAATTATCCTGGCTGTTAAGGCAATCATT 780

781 TTCAGACTGGACCCGGCTCTCATCCTGATCAGGAGCCCTATATCCTTACGTGGCAAGATT 840

841 TGGCAGAAGATCCTCCGCCATGGTTAAACCATGGCTAAATAACCAAGAAAGCCAGGTC 900

901 CCCGAATCCTGGCTTTGGAGAGAAAAACAAACACTCGGCCGAAAAAGTCGAGCCCTCTT 960

961 CCTCGTATCTACCCGAGATCGAGGAGCCGACTTGGCCGGAACCCCAACCTGTTCCC 1020

1021 CCACCCCCCTATCCAGCACAGGTGCTGTGAGGGGACCTCTGCCCTCTGGAGCTCCGG 1080

1081 TGGTGGAGGGACCTGCTGCCGGACTCGGAGCCGGAGAGGCCACCCGGAGCGGGACAG 1140

1141 ACGAGATCGCGATATTACCGCTGCGCACCTATGCCCTCCATGCCAGGGGCCATTGC 1200

1201 AGCCCTCCAGTATTGGCCCTTTCTCTGAGATCTCTATAATTGGAAAACTAACCATC 1260

5/22

Figure 2 cont.

1261	CCCCTTCTCGGAGGATCCCCAACGCCTCACGGGTTGGTGGAGTCCCTATGTTCTCTC	1320
1321	ACCAGCCTACTTGGGATGATTGTCAACAGCTGCTGCAGACACTCTTCACAACCGAGGAGC	1380
1381	GAGAGAGAATTCTGTTAGAGGCTAGAAAAATGTTCTGGGCCGACGGCGACCCACGC	1440
1441	AGTTGCAAAATGAGATTGACATGGGATTCCCTGACTCGCCCCGGTTGGGACTACAACA	1500
1501	CGGCTGAAGGTAGGGAGAGCTTGAATCTATGCCAGGCTCTGGTGGCGGGCTCCGGG	1560
1561	GCGCCTCAAGACGGCCCCTAAATTGGCTAAGGTAAGAGAGGGTGTGCAGGGACCGAACG	1620
1621	AACCTCCCTCGGTATTCCTGAGAGGCTCATGGAAGCCTTCAGGCCGTTCAACCCCTTTG	1680
1681	ATCCTACCTCAGAGGCCAGAAAGCCTCAGTGGCCCTGGCCTTCATTGGCAGTCGGCTC	1740
1741	TGGATATCAGGAAGAACTTCAGAGACTGGAAGGGTTACAGGAGGCTGAGTTACGTGATC	1800
1801	TAGTGAGAGAGGCAGAGAAGGTGTATTACAGAAGGGAGACAGAAGAGGAGAAGGAACAGA	1860
1861	GAAAAGAAAAGGGAGAGAGAAGAAAGGGAGGAAGACGTGATAGACGGCAAGAGAAAGAATT	1920
1921	TGACTAAGATCTTGGCCGCAGTGGTTGAAGGAAAGAGCAGCAGGGAGAGAGAGAGATT	1980
1981	TTAGGAAAATTAGGTCAAGGCCCTAGACAGTCAGGGAACCTGGCAATAGGACCCACTCG	2040
2041	ACAAGGACCAGTGTGCGTATTGTAAGAAAAGGACACTGGCAAGGAAGTGCCTTGGCAAGA	2100
2101	AGGGAAACAAAGGACCGAAGTCCTAGCTCTAGAAGAAGATAAAGATTAGGGAGACGGGT	2160
2161	TCGGACCCCTCCCCGAGCCAGGGTAACCTTGAAGGTGGAGGGCAACCAGTTGAGTTC	2220
2221	CTGGTTGATACCGGAGCGGAGCATTCACTGCTGCTACAACCATTAGGAAAAGTAAAGAA	2280
2281	AAAAAAATCCTGGGTGATGGGTGCCACAGGGCAACGGCAGTATCCATGGACTACCGAAGA	2340
2341	ACCGTTGACTTGGGAGTGGGACGGGTAACCCACTCGTTCTGGTCATCCCTGAGTGCCTA	2400
2401	GTACCCCTCTAGGTAGAGACTTACTGACCAAGATGGAGCTAAATTCTTTGAACAA	2460
2461	GGAAGACCAGAAGTGTCTGTGAATAACAAACCCATCACTGTGTTGACCCCTCCAATTAGAT	2520

6/22

Figure 2 cont.

2521	GATGAATATCGACTATATTCTCCCCAAGTAAAGCCTGATCAAGATATACTACAGTCCTGGTTG	2580
2581	GAGCAGTTCCCCAAGCCTGGCAGAAACCGCAGGGATGGGTTGGCAAAGCAAGTTCCC	2640
2641	CCACAGGTTATTCAACTGAAGGCCAGTGCTACACCACTATCAGTCAGACAGTACCCCTTG	2700
2701	AGTAGAGAGGCTCGAGAAGGAATTGGCCGCATGTTCAAAGATTAATCCAACAGGGCATC	2760
2761	CTAGTTCCGTCCAATCCCCTGGAAATACTCCCCTGCTACCGGTTAGGAAGCCTGGGACC	2820
2821	AATGATTATCGACCACTACAGGACTTGAGAGAGGTCAATAAAAGGGTGCAGGACATACAC	2880
2881	CCAACGGTCCCGAACCCCTTATAACCTCTTGAGCGCCCTCCCGCTGAACGGAACGGTAC	2940
2941	ACAGTATTGGACTTAAAGATGCCCTCTGAGATTACACCCACTAGCCAACCA	3000
3001	CTTTTGCCCTCGAATGGAGAGATCCAGGTACGGGAAGAACCGGGCAGCTCACCTGGACC	3060
3061	CGACTGCCCAAGGGTTCAAGAACTCCCCGACCATCTTGACGAAGCCCTACACAGGGAC	3120
3121	CTGGCCAACCTCAGGATCCAACACCCCTCAGGTGACCCCTCCAGTACGTGGATGACCTG	3180
3181	CTTCTGGCGGGAGCCAAACAGGACTGCTAGAAGGTACGAAGGCACTACTGCTGGAA	3240
3241	TTGTCTGACCTAGGCTACAGAGCCTCTGCTAAGAAGGCCAGATTGAGGAGAGGTA	3300
3301	ACATACTGGGTACAGTTGGGGGGGGCAGCGATGGCTGACGGAGGCACGGAAAGAAA	3360
3361	ACTGTAGTCCAGATACCGGCCCCAACACAGCCAAACAAGTGAGAGAGTTTGGGACA	3420
3421	GCTGGATTTGCAGACTGTGGATCCGGGTTGGCAGCTTAGCAGCCCCACTCTACCCG	3480
3481	CTAACCAAAGAAAAAGGGGGATTCTCCTGGCTCCTGAGCACCAGAAGCATTGATGCT	3540
3541	ATCAAAAAGGCCCTGCTGAGCGCACCTGCTCTGGCCCTCCCTGACGTAACAAACCTTT	3600
3601	ACCCTTATGTGGATGAGCGTAAGGGAGTAGCCCGAGGAGTTAACCCAAACCTAGGA	3660
3661	CCATGGAGGAGACCTGTTGCCTACCTGTCAAAGAAGCTTGTACCTGTAGCCAGTGGTTGG	3720
3721	CCCGTATGTGAAGGCTATCGCAGCTGTGGCCATACTGGTCAAGGACGCTGACAAATTG	3780

7/22

Figure 2 cont.

3781	ACTTTGGGACAGAATATAACTGTAATAGCCCCCATGCATTGGAGAACATCGTCGGCAG	3840
3841	CCCCCAGACCGATGGATGACCAACGCCGCATGACCCACTATCAAAGCCTGCTTCTCACA	3900
3901	GAGAGGGTCACTTCGCTCCACCAGCCGCTCTCAACCCCTGCCACTCTTCTGCCTGAAGAG	3960
3961	ACTGATGAACCAGTGACTCATGATTGCCATCAACTATTGATTGAGGAGACTGGGTCCGC	4020
4021	AAGGACCTTACAGACATAACCGCTGACTGGAGAACGTGCTAACCTGGTTACTGACGGAAGC	4080
4081	AGCTATGTGGTGGAAAGGTAAGAGGGATGGCTGGGGGGCAGTGGTGGACGGGACCCGCACG	4140
4141	ATCTGGGCCAGCAGCCTGCCCGGAAGGAACCTCAGCGAAAAGGCTGAGCTCATGGCCCTC	4200
4201	ACGCAAGCTTGGCGCTGGCGAAGGGAAATCCATAAACATTATACGGACAGCAGGTAT	4260
4261	GCCTTGCAGTGCACACGTACACGGGCCATCTATAAACAAAGGGGTTGCTTACCTCA	4320
4321	GCAGGGAGGGAAATAAGAACAAAGAGGGAAATTCTAACGCTATTAGAACGCTTACATTG	4380
4381	CCAAAAAAGGCTAGCTATTATACACTGTCCTGGACATCAGAAAGCCAAAGATCTCATATCT	4440
4441	AGAGGGAAACCAGATGGCTGACCGGGTTGCCAAGCAGGCAGCCCAGGCTGTTAACCTCTG	4500
4501	CCTATAATAGAAACGCCAAGCCCCAGAACCCAGACGACAGTACACCTAGAACAGACTGG	4560
4561	CAAGAGATAAAAAGATAGACCAAGTCTCTGAGACTCCGGAGGGGACCTGCTATAACCTCA	4620
4621	TATGGGAAGGAAATCTGCCCAACAAAGAAGGGTTAGAATATGTCCAACAGATACTCGT	4680
4681	CTAACCCACCTAGGAACCTAACACCTGCAGCAGTGGTCAGAACATCCCCTATCATGTT	4740
4741	CTGAGGCTACCAGGAGTGGCTGACTCGGTGGTCAAACATTGTGTGCCCTGCCAGCTGGTT	4800
4801	AATGCTAATCCTCCAGAAATACCTCCAGGAAAGAGAGACTAAGGGGAAGCCACCCAGGCGCT	4860
4861	CACTGGGAAGTGGACTTCACTGAGGTAAAGCCGGCTAAATACGGAAACAAATATCTATTG	4920
4921	GTTTTGTAGACACCTTTCAGGATGGTAGAGGGTTATCCTACTAAGAAAGAGACTTCA	4980
4981	ACCGTGGTGGCTAAGAAAATCTGGAGGAATTTCAGGATGGTAGAGGGTTATCCTACTAAGAAAGAGACTTCA	5040

8/22

Figure 2 cont.

5041 ATAGGGTCAGACAATGGTCCAGCTTCGTTGCCAGGTAAGTCAGGGACTGGCCAAGATA 5100

5101 TTGGGGATTGATTGGAAACTGCATTGTGCATAACAGACCCAAAGCTCAGGACAGGTAGAG 5160

5161 AGGATGAATAGAACCATTAAGAGAGACCCCTACCAAATTGACCACAGAGACTGGCATTAAAT 5220

5221 GATTGGATGGCTCTCCTGCCCTTGTGCTTTAGGGTGAGGAACACCCCTGGACAGTTT 5280

5281 GGGCTGACCCCTATGAATTGCTCTACGGGGACCCCCCCCCTGGCAGAAATTGCCCTT 5340

5341 GCACATAGTGTGATGTGCTGCTTCCCAGCCTTGTCTAGGCTCAAGGCCTCGAG 5400

5401 TGGGTGAGGCAGCGAGCGTGGAAAGCAGCTCCGGAGGCCTACTCAGGAGGAGACTTGCAA 5460

5461 GTTCCACATCGCTTCCAAGTTGGAGATTCACTATGTTAGACGCCACCGTGCAGGAAAC 5520

5521 CTCGAGACTCGGTGGAAGGGACCTTATCTCGTACTTTGACCACACCAACGGCTGTGAAA 5580

5581 GTCGAAGGAATCCCCACCTGGATCCATGCATCCCACGTTAAGCYGGCGCCACCTCCCGAC 5640

5641 TCGGGGTGGAGAGCCGAAAAGACGTGAGAATCCCCTAAGCTTGCCTCCATCGCCTGGTT 5700

5701 CCTTACTCTAACAAATAACTCCCCAGGCCAGTAGTAAACGCCCTATAGACAGCTCGAACCC 5760

5761 CCATAGACCTTTATCCCCACCTGGCTGATTATTGACCCCTGATAACGGGTGTCACTGTAAA 5820

5821 TAGCACTCGAGGTGTTGCTCTAGAGGCACCTGGTGGCCTGAACCTGCATTTCTGCCTCCG 5880

5881 ATTGATTAACCCCGCTGTTAARAGCACACCTCCAACCTAGTCCGTAGTTATGGGTTCTA 5940

5941 TTGCTGCCAGGCACAGAGAAAGAGAAATACTGTGGGGTTCTGGGAATCCTTCTGTAG 6000

6001 GAGATGGAGCTCGTCAACCTCCAACGATGGAGACTGGAAATGCCGATCTCTCCAGGA 6060

6061 CCGGGTAAAATTCTCCTTGTCAATTCCGGCCGGCAAGTACAAAATGATGAAACTATA 6120

6121 TAAAGATAAGAGCTGCTCCCCATCAGACTTAGATTATCTAAAGATAAGTTCACTGAAAG 6180

6181 GAAAACAGGAAAATATTCAAAAGTGGATAAAATGGTATGAGCTGGGAATAGTTTTTATT 6240

6241 ATATGGCGGGGGAGCAGGGTCCACTTTAACCAATTGCCCTTAGGATAGAGACGGGGACAGA 6300

9/22

Figure 2 cont.

6301	ACCCCTGTGGCAATGGGACCGATAAAAGTACTGGCTGAACAGGGGCCCCGGCCCTGGA	6360
6361	GCCACCGCATAACTGCCGGTGCCCCAATTAAACCTCGCTGCCGCTGACATAACACAGCC	6420
6421	GCCTAGCAACAGTACCACTGGATTGATTCTACCAACACGCCAGAAACTCCCCAGGTGT	6480
6481	TCCTGTTAAGACAGGGACAGAGACTCTTCAGTCTCATCCAGGGAGCTTCCAAGCCATCAA	6540
6541	CTCCACCGACCCGTATGCCACTTCTTGTGGCTTGTCTATCCTCAGGGCCTCTTA	6600
6601	TTATGAGGGGATGGCTAAAGAAAAGAAAATTCAATGTGACCAAAGAGGCATAGAAATCAATG	6660
6661	TACATGGGGTCCCGAAATAAGCTTACCCCTCACTGAAGTTCCGGAAAGGGGACATGCCAT	6720
6721	AGGAAAAGCTCCCCATCCCACCAACACCTTGCTATAGTACTGTGGTTATGAGCAGGC	6780
6781	CTCAGAAAATCAGTATTTAGTACCTGGTTATAACAGGTGGTGGCATGCAATACTGGTT	6840
6841	AACCCCTGTGTTCCACCTCAGTCTCAACCAATCCAAAGATTCTGTGTATGGTCCA	6900
6901	AATCGTCCCCCGAGTGTACTACCATCCTGAGGAAGTGGTCCTGATGAATAATGACTATCG	6960
6961	GTATAACCGACCAAAAGAGAACCCGTATCCCTACCCCTAGCTGTAATGCTCGGATTAGG	7020
7021	GACGGCCGTTGGCGTAGGAACAGGGACAGCTGCCCTGATCACAGGACCACAGCAGCTAGA	7080
7081	GAAAGGACTTGGTGAGCTACATGCGGCCATGACAGAAGATCTCCGAGCCTAAAGGAGTC	7140
7141	TGTTAGCAACCTAGAAGAGTCCCTGACTTCTTGTCTGAAGTGGTTCTACAGAACCGGAG	7200
7201	GGGATTAGATCTGCTGTTCTAAGAGAAGGTGGTTATGTGCAGCCTAAAAGAAGAATG	7260
7261	TTGCTTCTATGTAGATCACTCAGGAGCCATCAGAGACTCCATGAACAAGCTTAGAAAAAA	7320
7321	GTTAGAGAGGCGTCGAAGGGAAAAGAGAGGGCTGACCAGGGTGGTTGAAGGATGGTTCAA	7380
7381	CAGGTCTCCTGGATGACCACCCGTCTTCTGCTCTGACGGGCCCCTAGTAGTCCTGCT	7440
7441	CCTGTTACTTACAGTTGGCCTTGCTTAATTAAATAGGTTGTCCTTGTAGAGAACG	7500
7501	AGTGAGTGCAGTCCAGATCATGGTACTTAGGCAACAGTACCAAGGCCTCTGAGCCAAGG	7560

10/22

Figure 2 cont.

7561 AGAAAATGACCTCTAGCCTTCCCAGTTCTAAGATTAGAACTATTAACAAGACAAGAAGTG 7620
7621 GGGAAATGAAAGGATGAAAATGCAACCTAACCCCTCCCAGAACCCAGGAAGTTAATAAAAAG 7680
7681 CTCTAAATGCCCGAATTMCAGACCCCTGCTGGCTGCCAGTAAATAGGTAGAAGGTCACA 7740
7741 CTTCCCTATTGTTCCAGGGCCTGCTATCCTGGCTAAGTAAGATAACAGGAAATGAGTTGA 7800
7801 CTAATCGCTTATCTGGATTCTGTAAAATGACTGGCACCATAGAAGAATTGATTACACAT 7860
7861 TGACAGCCCTAGTGACCTATCTCAACTGCAATCTGTCACTCTGCCAGGAGGCCACGCAG 7920
7921 ATGCGGACCTCCGGAGCTATTTAAAATGATTGGTCCACGGAGCGCGGGCTCTCGATATT 7980
7981 TTAAAATGATTGGTCCATGGAGCGCGGGCTCTCGATATTTAAAATGATTGGTTGTGAC 8040
8041 GCACAGGCTTGTGAACCCATAAAAGCTGTCCCGATTCCGCACTCGGGGCCGCAGT 8100
8101 CCTCTACCCCTGCGTGGTGTACGACTGTGGGCCAGCGCGCTTGGAAATAAAATCCTCT 8160
8161 TGCTGTTGCATAAAAAAAAAAAAAAAAAAAAA 8196

Figure 3

11/22

1 GTGGTGTACGACTGTGGGCCAGCGCGCTTGAATAAAATCCTCTGCTGTTGCATC 60

61 AAGACCGCTTCTCGTGAGTGATTGGGGTGTGCCTCTTCCGAGCCGGACGAGGGGAT 120

121 TGTTCTTTACTGGCCTTCATTGGTGCCTGGCCGGAAATCCTGCGACCACCCCTTA 180

181 CACCCGAGAACCGACTTGGAGGTAAAGGGATCCCCTTGGAACGTGTGTGTGGCC 240

241 GGC GTCTCTGTTCTGAGTGCTGTTTGGTGTGCCTTCGGTTGCAGCTGTCCT 300

301 CTCAGACCGTAAGGACTGGAGGACTGTGATCAGCAGACGTGCTAGGAGGATCACAGGCTG 360

361 CCACCCCTGGGGACGCCCGGGAGGTGGGAGAGCCAGGGACGCCGGTGGTCTCTACT 420

421 GTCGGTCAGAGGACCGAGTTCTGTTGAAGCGAAAGCTCCCCCTCCGGCCGTCCG 480

481 ACTCTTTGCCTGCTTGTGGAAGACGCCGGACGGTGCCTGCTGGATCTGTTGGTT 540

541 CTGTCTCGTGTCTTGTCTTGTACAGTTAAATATGGACAGACAG
MetGlyGlnThrV 600

601 TGACTACCCCCCTTAGTTGACTCTGACCATGGACTGAAGTTAGATCCAGGGCTCATA
alThsThsProLeuSerLeuThrLeuAspHisTrpThrGluValArgSerArgAlaHisA 660

661 ATTTGTCAAGTTCAAGAAGGGACCTTGGCAGACTTCTGTGCCTCTGAATGCCAA
snLeuSerValGlnValLysLysGlyProTrpGlnThrPheCysAlaSerGluTrpProT 720

721 CATTGATGTTGGATGCCATCAGAGGGACCTTAATTCTGAAATTATCCTGGCTGTTA
hrPheAspValGlyTrpProSerGluGlyThrPheAsnSerGluIleIleLeuAlaValL 780

781 AGGCAATCATTTCAGACTGGACCCGGCTCTCATCCTGATCAGGAGCCCTATATCCTTA
ysAlaIleIlePheGlnThrGlyProGlySerHisProAspGlnGluProTyrIleLeuT 840

841 CGTGGCAAGATTGGCAGAAGATCCTCCGCCATGGTTAAACCATGGCTAAATAACCAA
hrTrpGlnAspLeuAlaGluAspProProProTrpValLysProTrpLeuAsnLysProA 900

901 GAAAGCCAGGTCCCCGAATCCTGGCTCTGGAGAGAAAAACAAACACTGGCCGAAAAAG
rgLysProGlyProArgIleLeuAlaLeuGlyGluLysAsnLysHisSerAlaGluLysV 960

961 TCGAGCCCTCTCCTCGTATCTACCCCGAGATCGAGGGAGCCGCCGACTTGGCCGGAACCC
alGluProSerSerSerTyrLeuProArgAspArgGlyAlaAlaAspLeuAlaGlyThrP 1020

1021 CAACCTGTTCCCCACCCCTTATCCAGCACAGGGTGCTGTGAGGGGACCTCTGCCCTC
roThrCysSerProThrProLeuSerSerThrGlyCysCysGluGlyThrSerAlaProP 1080

12/22

Figure 3 cont.

1081 CTGGAGCTCCGGTGGAGGGACCTGCTGCCGGACTCGGAGGCCGGAGAGGCCACCC 1140
roGlyAlaProValValGluGlyProAlaAlaGlyThrArgSerArgArgGlyAlaThrP

1141 CGGAGCGGACAGACGAGATCGCGATATTACCGCTGCGCACCTATGGCCCTCCATGCCAG 1200
roGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyrGlyProProMetProG

1201 GGGGCCAATTGCAGCCCCCTCCAGTATTGGCCCTTTCTTCTGCAGATCTCTATAATTGGA 1260
lyGlyGlnLeuGlnProLeuGlnTyrTrpProPheSerSerAlaAspLeuTyrAsnTspL

1261 AAACTAACCATCCCCCTTCGGAGGATCCCCAACGCCTCACGGGGTTGGTGGAGTCCC 1320
ysThrAsnHisProProPheSerGluAspProGlnArgLeuThrGlyLeuValGluSerL

1321 TTATGTTCTCTCACCGCCTACTTGGGATGATTGTCAACAGCTGCTGCAGACACTCTTCR 1380
euMetPheSerHisGlnProThrTrpAspAspCysGlnGlnLeuLeuGlnThrLeuPheT

1381 CAACCGAGGAGCGAGAGAGAATTCTGTTAGAGGCTAGAAAAAAATGTTCTGGGGCCGACG 1440
hrThrGluGluArgGluArgIleLeuLeuGluAlaArgLysAsnValProGlyAlaAspG

1441 GGCGACCCACGCAGTTGCAAAATGAGATTGACATGGGATTCCCTTGACTCGCCCCGGTT 1500
lyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProLeuThrArgProGlyT

1501 GGGACTACAACACGGCTGAAGGTAGGGAGAGCTTGAAATCTATGCCAGGCTCTGGTGG 1560
rPAspTyrAsnThrAlaGluGlyArgGluSerLeuLysIleTyrArgGlnAlaLeuValA

1561 CGGGTCTCGGGGGCGCCTCAAGACGGCCCCTAATTGGCTAAGGTAAGAGAGGGTGTGC 1620
laGlyLeuArgGlyAlaSerArgArgProThrAsnLeuAlaLysValArgGluValMetG

1621 AGGGACCGAACGAACTCCCTCGGTATTCTTGAGAGGCTCATGGAAGCCTTCAGGCAGG 1680
InGlyProAsnGluProProSerValPheLeuGluArgLeuMetGluAlaPheArgArg?

1681 TCACCCCTTGATCCTACCTCAGAGGCCAGAAAGCCTCAGTGGCCCTGGCCTTCATTG 1740
heThrProPheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPheIleG

1741 GGCAGTCGGCTCTGGATATCAGGAAGAAACTTCAGAGACTGGAAGGGTTACAGGAGGCTG 1800
lyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGluGlyLeuGlnGluAlaG

1801 AGTTACGTGATCTAGTGAGAGAGGCAGAGAAGGTGTATTACAGAAGGGAGACAGAAGAGG 1860
luLeuArgAspLeuValArgGluAlaGluLysValTyrTyrArgArgGluThrGluGluG

1861 AGAAGGAACAGAGAAAAGAAAAGGAGAGAGAAGAAAGGGAGGAAAGACGTGATAGACGGC 1920
luLysGluGlnArgLysGluLysGluArgGluGluArgGluArgArgAspArgArgG

1921 AAGAGAAGAATTGACTAACGATCTTGGCCGCAGTGGTTGAAGGGAAGAGCAGCAGGGAGA 1980
InGluLysAsnLeuThrLysIleLeuAlaAlaValValGluGlyLysSerSerArgGluA

1981 GAGAGAGAGATTAGGAAAATTAGGTAGGCCCTAGACAGTCAGGGAACCTGGCAATA 2040
rgGluArgAspPheArgLysIleArgSerGlyProArgGlnSerGlyAsnLeuGlyAsnA

Figure 3 cont

2041 GGACCCCACTCGACAAGGACCAGTGTGCGTATTGTAAAGAAAAAGGACACTGGGCAAGGA 2100
 rgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGlyHisTrpAlaArgA

2101 ACTGCCCAAGAAGGGAAACAAAGGACCGAAGgTCCTAGCTCTAGAAGAAGATAAAGATT 2160
 snCysProLysLysGlyAsnLysGlyProLysValLeuAlaLeuGluGluAspLysAspE

2161 AGGGGAGACGGGgTTCGGACCCCCTCCCCGAGCCCAGGGTAACTTGAAGGTGGAGGGC 2220
 ndGlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysValGluGlyG

2221 AACCAAGTTGAGTTCTGGTTGATACCGGAGCGGAGCATTCACTGCTGCTACAACCATTAG 2280
 lnProValGluPheLeuValAspThrGlyAlaGluHisSerValLeuLeuGlnProLeuG

2281 GAAAACTAAAAGAAAAAAATCCTGGGTGATGGGTGCCACAGGGCAACGGCAGTATCCAT 2340
 lyLysLeuLysGluLysLysSerTrpValMetGlyAlaThrGlyGlnArgGlnTyrProT

2341 GGACTACCCGAAGAACCGTTGACTTGGGAGTGGGACGGGTAACCCACTCGTTCTGGTCA 2400
 rpThrThrArgArgThrValAspLeuGlyValGlyArgValThrHisSerPheLeuValI

2401 TCCCTGAGTGCeAGTACCCCTCTAGGTAGAGACTTACTGACCAAGATGGGAGCTCAAA 2460
 leProGluCysProValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnI

2461 TTTCTTTGAACAAGGAAGACCAGAAGTGTCTGTGAATAACAAACCCATCACTGTGTTGA 2520
 leSerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleThrValLeut

2521 CCCTCCAATTAGATGATGAATATCGACTATATTCTCCCCAAGTAAAGCCTGATCAAGATA 2580
 hrLeuGlnLeuAspAspGluTyrSerProGlnValLysProAspGlnAspI

2581 TACAGTCTGGTGGAGCAGTTCCCCAACGCTGGCAGAAACCGCAGGGATGGGTTGG 2640
 leGlnSerTrpLeuGluGlnPheProGlnAlaTrpAlaGluThrAlaGlyMetGlyLeuA

2641 CAAAGCAAGTTCCCCACAGGTTATTCAACTGAAGGCCAGTGTCTACACCAGTATCAGTC 2700
 laLysGlnValProProGlnValIleGlnLeuLysAlaSerAlaThrProValSerValA

2701 GACAGTACCCCTTGAGTAGAGAGGGCTCGAGAAGGAATTGGCCGCATGTTCAAAGATTAA 2760
 rgGlnTyrProLeuSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuI

2761 TCCAACAGGGCATCTAGTTCTGTCCAATCCCCCTGGAATACTCCCCTGCTACCGGTTA 2820
 leGlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLeuProValA

2821 GGAAGCCTGGGACCAATGATTATCGACCAAGTACAGGACTTGAGAGAGGGTCAATAAAAGGG 2880
 rgLysProGlyThrAsnAspTyrArgProValGlnAspLeuArgGluValAsnLysArgV

2881 TGCAGGACATACACCCAACGGTCCCGAACCCCTTATAACCTCTTGAGCGCCCTCCGCCTG 2940
 alGlnAspIleHisProThrValProAsnProTyrAsnLeuLeuSerAlaLeuProProG

2941 AACGGAACCTGGTACACAGTATTGGACTAAAAGATGCCTTCTTCTGCCTGAGATTACACC 3000
 luArgAsnTrpTyrThrValLeuAspLeuLysAspAlaPhePheCysLeuArgLeuHisP

Figure 3 cont.

3001 CCACTAGCCAACCACCTTTGCCTTCGAATGGAGAGATCCAGGTACGGGAAGAACCGGGC 3060
 roThrSerGlnProLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyG

3061 AGCTCACCTGGACCCGACTGCCCAAGGGTTCAAGAACTCCCCGACCATCTTGACGAAG 3120
 lnLeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePheAspGluA

3121 CCCTACACAGGGACCTGGCCAACCTCAGGATCCAACACCCCTCAGGTGACCCCTCCAGT 3180
 laLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProGlnValThrLeuLeuGlnT

3181 ACGTGGATGACCTGCTTCTGGCGGGAGCCACCAAACAGGACTGCTTAGAAGGTACGAAGG 3240
 yrValAspAspLeuLeuAlaGlyAlaThrLysGlnAspCysLeuGluGlyThrLysA

3241 CACTACTGCTGGAATTGTCTGACCTAGGCTACAGAGCCTCTGCTAAGAAGGCCAGATT 3300
 laLeuLeuLeuGluLeuSerAspLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleC

3301 GCAGGAGAGAGGTAACATACTGGGGTACAGTTGCAGGGGGGGCAGCGATGGCTGACGG 3360
 ysArgArgGluValThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrG

3361 AGGCACGGAAGAAAAGTGTAGTCCAGATAACGGCCCCAACACAGCCAAACAAGTGAGAG 3420
 luAlaArgLysLysThrValValGinIleProAlaProThrAlaLysGlnValAlaG

3421 AGTTTTGGGACAGCTGGATTTGCAGACTGTGGATCCCAGGGTTTGCACCTTAGCAG 3480
 luPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProGlyPheAlaThrLeuAlaA

3481 CCCCACACTACCCGCTAACCAAAGAAAAAGGGGGATTCTCCTGGCTCCTGAGCACCAAGA 3540
 laProLeuTyrProLeuThrLysGluLysGlyGlyPheSerTrpAlaProGluHisGlnL

3541 AGGCATTGATGCTATCAAAAGGCCCTGCTGAGCGCACCTGCTCTGGCCCTCCCTGACG 3600
 ysAlaPheAspAlaIleLysLysAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspV

3601 TAACTAAACCCCTTACCCCTTATGTGGATGAGCGTAAGGGAGTAGCCCGAGGAGTTTAA 3660
 alThrLysProPheThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuT

3661 CCCAAACCCCTAGGACCACGGAGGGAGACCTGTTGCCTACCTGTCAAAGAAGCTTGATCCTG 3720
 hrGlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLeuAspProV

3721 TAGCCAGTGGTTGGCCCGTATGTCTGAAGGCTATCGCAGCTGTGGCCATACTGGTCAAGG 3780
 alAlaSerGlyTrpProValCysLeuLysAlaIleAlaAlaValAlaIleLeuValLysA

3781 ACGCTGACAAATTGACTTTGGGACAGAAATATAACTGTAATAGCCCCCATGCATTGGAGA 3840
 spAlaAspLysLeuThrLeuGlyGlnAsnIleThrValIleAlaProHisAlaLeuGlnA

3841 ACATCGTTGGCAGCCCCCAGACCGATGGATGACCAACGCCGCATGACCCACTATCAA 3900
 snIleValArgGlnProProAspArgTrpMetThrAsnAlaArgMetThrHisTyrGlnS

3901 GCCTGCTTCTCACAGAGAGGGTCACTTTCGCTCCACAGCCGCTCTCAACCCCTGCCACTC 3960
 erLeuLeuLeuThrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrL

Figure 3 cont

3961 TTCTGCCCTGAAGAGACTGATGAACCAGTGACTCATGATTGCCATCAACTATTGATTGAGG 4020
euLeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLeuIleGluG

4021 AGACTGGGGTCCGCAAGGACCTTACAGACATACCGCTGACTGGAGAAGTGCTAACCTGGT 4080
luThrGlyValArgLysAspLeuThrAspIleProLeuThrGlyGluValLeuThrTrpP

4081 TCACTGACGGAAGCAGCTATGTGGTGGAAAGGTAAGAGGGATGGCTGGGGCGGCAGTGGTGG 4140
heThrAspGlySerSerTyrValValGluGlyLysArgMetAlaGlyAlaAlaValValA

4141 ACAGGGACCCGCACGATCTGGGCCAGCAGCCTGCCGGAAAGGAACCTCAGCGCAAAAGGCTG 4200
spGlyThrArgThrIleTrpAlaSerSerLeuProGluGlyThrSerAlaGlnLysAlaG

4201 AGCTCATGGCCCTCACGCAAGCTTGGCGCTGGCCGAAGGGAAATCCATAAACATTTATA 4260
luLeuMetAlaLeuThrGinAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrT

4261 CGGACAGCAGGTATGCCCTTGCAGACTGCACACGTACACGGGCCATCTATAAACAAAGGG 4320
hrAspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLysGlnArgG

4321 GGTTGCTTACCTCAGCAGGGAGGGAAATAAGAACAAAGAGGGAAATTCTAACGCTATTAG 4380
lyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluGluIleLeuSerLeuLeuG

4381 AAGCCTTACATTGCCAAAAAGGCTAGCTATTACACTGTCTGGACATCAGAAAGCCA 4440
luAlaLeuHisLeuProLysArgLeuAlaIleIleHisCysProGlyHisGlnLysAlaL

4441 AAGATCTCATATCTAGAGGGAAACCAGATGGCTGACCGGGTTGCCAACGCAGGCCAGG 4500
ysAspLeuIleSerArgGlyAsnGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnA

4501 CTGTTAACCTCTGCCCTATAATAGAAACGCCAAAGCCCCAGAACCCAGACGACAGTACA 4560
laValAsnLeuLeuProIleIleGluThrProLysAlaProGluProArgArgGlnTyrT

4561 CCCTAGAAGACTGCCAAGAGATAAAAAGATAGACCAGTTCTCTGAGACTCCGGAGGGGA 4620
hrLeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrProGluGlyT

4621 CCTGCTATACCTCATATGGGAAGGAAATCCTGCCCCACAAAGAACGGTTAGAATATGTCC 4680
hrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysGluGlyLeuGluTyrValG

4681 AACAGATAACATCGCTAACCCACCTAGGAACCTAACACACCTGCAGCAGTTGGTCAGAACAT 4740
lnGlnIleHisArgLeuThrHisLeuGlyThrLysHisLeuGlnGlnLeuValArgThrS

4741 CCCCTTATCATGTTCTGAGGCTACCAAGGAGTGGCTGACTCGGTGGTCAAACATTGTGTGC 4800
erProTyrHisValLeuArgLeuProGlyValAlaAspSerValValLysHisCysValP

4801 CCTGCCAGCTGGTTATGCTAATCCTTCCAGAAATACCTCCAGGAAAGAGAGACTAAGGGGAA 4860
roCysGlnLeuValAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlyS

4861 GCCACCCAGGCGCTCACTGGGAAGTGGACTTCAGTGAGGTAAAGCCGGCTAAATACGGAA 4920
erHisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLysTyrGlyA

Figure 3 cont.

4921 ACAAAATATCTATTGGTTTTGTAGACACCTTTCAGGATGGGTAGAGGCTTATCCTACTA 4980
snLysTyrLeuLeuValPheValAspThrPheSerGlyTrpValGluAlaTyrProThrL

4981 AGAAAAGAGACTTCAACCGTGGTGGCTAAGAAAATACTGGAGGAAATTTTCCAAGATTG 5040
ysLysGluThrSerThrValValAlaLysLysIleLeuGluGluIlePheProArgPheG

5041 GAATACCTAAGGTAAAGGGTCAGACAATGGTCCAGCTTCGTTGCCAGGTAAAGTCAGG 5100
lyIleProLysValIleGlySerAspAsnGlyProAlaPheValAlaGlnValSerGlnG

5101 GACTGGCCAAGATATTGGGGATTGGATTGGAAACTGCATTGTGCATACAGACCCCCAAAGCT 5160
lyLeuAlaLysIleLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerS

5161 CAGGACAGGTAGAGAGGGATGAAATAGAACCATTAAGAGACCCCTTACCAAATTGACCACAG 5220
erGlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLeuThrThrG

5221 AGACTGGCATTAAATGATTGGATGGCTCTCCCTGCCCTTGTGCTTTAGGGTGAGGAACA 5280
luThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValLeuPheArgValArgAsnT

5281 CCCCTGGACAGTTGGCTGACCCCCCTATGAATTGCTCTACGGGGGACCCCCCCCCTTGG 5340
hrProGlyGlnPheGlyLeuThrProTyrGluLeuLeuTyrGlyGlyProProProLeuA

5341 CAGAAATTGCCTTGCACATAGTGCTGATGTGCTGCTTCCCAGCCTTGTCTCTAGGC 5400
laGluIleAlaPheAlaHisSerAlaAspValLeuLeuSerGlnProLeuPheSerArgL

5401 TCAAGGGCGCTCGAGTGGGTGAGGCAGCGAGCGTGGAAAGCAGCTCCGGGAGGCCTACTCAG 5460
euLysAlaLeuGluTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSerG

5461 GAGGAGACTTGCAGATTCCACATCGCTTCAAGTTGGAGATTCACTCTATGTTAGACGCC 5520
lyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTyrValArgArgH

5521 ACCGTGCAGGAAACCTCGAGACTCGGTGGAAGGGACCTTATCTCGTACTTTGACCACAC 5580
isArgAlaGlyAsnLeuGluThrArgTrpLysGlyProTyrLeuValLeuLeuThrThrP

5581 CAACGGCTGTGAAAGTCGAAGGAATCCCCACCTGGATCCATGCATCCPACGTTAACGCCGG 5640
roThrAlaValLysValGluGlyIleProThrTrpIleHisAlaSerHisValLysProA
MetHisProThrLeuSerArg

5641 CGCCACCTCCCGACTCGGGTGGAGAGCCGAAAAGActTGAGAATCCCCCTAACGCTTCGCC 5700
laProProProAspSerGlyTrpArgAlaGluLysThrGluAsnProLeuLysLeuArgL
ArgHisLeuProThrArgGlyGlyGluProLysArgLeuArgIleProLeuSerPheAla

5701 TCCATCGCCTGGTTCTTACTCTAACAAACTCCCCAGGCCAGTAGTAAACGCCCTTATA 5760
euHisArgLeuValProTyrSerAsnAsnAsnSerProGlyGlnEnd
SerIleAlaTrpPheLeuThrLeuThrIleThrProGlnAlaSerSerLysArgLeuIle

5761 GACAGCTCGAACCCCCATAGACCTTATCCCTTACCTGGCTGATTATTGACCCCTGATACG 5820
AspSerSerAsnProHisArgProLeuSerLeuThrTrpLeuIleIleAspProAspThr

17/22

Figure 3 cont

5821 GGTGTCAGTGTAAATAGCACTCGAGGTGTTGCTCCTAGAGGCACCTGGTGGCCTGAACTG 5880
 GlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrpTrpProGluLeu

5881 CATTTCCTGCCTCCGATTGATTAACCCCGCTGTTAAAAGCACACCTCCAACCTAGTCCGT 5940
 HisPheCysLeuArgLeuIleAsnProAlaValLysSerThrProProAsnLeuValArg

5941 AGTTATGGGTTCTATTGCTGCCAGGCACAGAGAAAGAGAAATACTGTGGGGTTCTGGG 6000
 SerTyrGlyPheTyrCysCysProGlyThrGluLysGluLysTyrCysGlyGlySerGly

6001 GAATCCTTCTGTAGGAGATGGAGCTGCGTCACCTCCAACGATGGAGACTGGAAATGGCCG 6060
 GluSerPheCysArgArgTrpSerCysValThrSerAsnAspGlyAspTrpLysTrpPro

6061 ATCTCTCTCCAGGACCGGGTAAAATTCTCCTTGTCATTCCGGCCCGGGCAAGTACAAA 6120
 IleSerLeuGlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrLys

6121 ATGATGAAAATATAAGATAAGAGCTGCTCCCCATCAGACTTAGATTATCTAAAGATA 6180
 MetMetLysLeuTyrLysAspLysSerCysSerProSerAspLeuAspTyrLeuLysIle

6181 AGTTTCACTGAAAGGAAAACAGGAAAATATTCAAAGTGGATAAAATGGTATGAGCTGGG 6240
 SerPheThrGluArgLysThrGlyLysTyrSerLysValAspLysTrpTyrGluLeuGly

6241 AATAGTTTTTATTATATGGCGGGGGAGCAGGGTCCACTTAACCATTGGCCTTAGGATA 6300
 AsnSerPheLeuLeuTyrGlyGlyAlaGlySerThrLeuThrIleArgLeuArgIle

6301 GAGACGGGGACAGAACCCCTGTGGCAATGGGACCCGATAAAAGTACTGGCTGAACAGGG 6360
 GluThrGlyThrGluProProValAlaMetGlyProAspLysValLeuAlaGluGlnGly

6361 CCCCCGGCCCTGGAGCCACCGCATAACTTGCCGGTGCCTAACCTCGCTGCGGCCT 6420
 ProProAlaLeuGluProProHisAsnLeuProValPheGlnLeuThrSerLeuArgPro

6421 GACATAACACAGCCGCCTAGCAGACAGTACCACTGGATTGATTCTACCAACACGCCTAGA 6480
 AspIleThrGlnProProSerAsnSerThrThrGlyLeuIleProThrAsnThrProArg

6481 AACTCCCCAGGTGTTCTGTTAAGACAGGACAGAGACTCTCAGTCTCATCCAGGGAGCT 6540
 AsnSerProGlyValProValLysThrGlyGlnArgLeuPheSerLeuIleGlnGlyAla

6541 TTCCAAGCCATCAACTCCACCGACCCCTGATGCCACTTCTTGTGGCTTGTCTATCC 6600
 PheGlnAlaIleAsnSerThrAspProAspAlaThrSerSerCysTrpLeuCysLeuSer

6601 TCAGGGCCTCCTTATTATGAGGGGATGGCTAAAGAAAGAAAATTCAATGTGACCAAAGAG 6660
 SerGlyProProTyrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGlu

6661 CATAGAAAATCAATGTACATGGGGGTCCGAAATAAGCTTACCCACTGAAGTTCCGGG 6720
 HisArgAsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSerGly

6721 AAGGGGACATGCATAGGAAAAGCTCCCCATCCCACCAACACCTTGCTATAGTACTGTG 6780
 LysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeuCysTyrSerThrVal

Figure 3 cont.

18/22

6781	GTTCATGAGCAGGCCCTCAGAAAATCAGTATTTAGTACCTGGTTATAACAGGTGGTGGGCA ValTyrGluGlnAlaSerGluAsnGlnTyrLeuValProGlyTyrAsnArgTrpTrpAla	6840
6841	TGCAATACTGGGTTAACCCCTGTGTTCCACCTCAGTCTCAACCAATCCAAAGATTTC CysAsnThrGlyLeuThrProCysValSerThrSerValPheAsnGlnSerLysAspPhe	6900
6901	TGTGTCATGGTCCAAATCGTCCCCCGAGTGTACTACCATCCTGAGGAAGTGGTCCTTGAT CysValMetValGlnIleValProArgValTyrTyrHisProGluGluValValLeuAsp	6960
6961	GAATATGACTATCGGTATAACCGACCAAAAGAGAACCCGTATCCCTTACCCTAGCTGTA GluTyrAspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAlaVal	7020
7021	ATGCTCGGATTAGGGACGGCGTTGGCGTAGGAAACAGGGACAGCTGCCCTGATCACAGGA MetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAlaAlaLeuIleThrGly	7080
7081	CCACAGCAGCTAGAGAAAGGACTTGGTGAGCTACATGCGGCCATGACAGAAAGATCTCCGA ProGlnGlnLeuGluLysGlyLeuGlyGluLeuHisAlaAlaMetThrGluAspLeuArg	7140
7141	GCCTTAAAGGAGTCTGTTAGCAACCTAGAAGAGTCCCTGACTTCTTGCTGAAGTGGTT AlaLeuLysGluSerValSerAsnLeuGluSerLeuThrSerLeuSerGluValVal	7200
7201	CTACAGAACCGGAGGGGATTAGATCTGCTGTTCTAAGAGAACGGTGGTTATGTGCAGCC LeuGlnAsnArgArgGlyLeuAspLeuLeuPheLeuArgGluGlyLeuCysAlaAla	7260
7261	TTAAAAGAAGAATGTTGCTTCTATGTAGATCACTCAGGAGGCCATCAGAGACTCCATGAAC LeuLysGluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMetAsn	7320
7321	AAGCTTAGAAAAAAAGTTAGAGAGGGCGTCGAAGGGAAAGAGAGGGCTGACCAGGGTGGTT LysLeuArgLysLysLeuGluArgArgArgGluArgAlaAspGlnGlyTyrPhe	7380
7381	GAAGGATGGTTAACAGGTCTCCTGGATGACCACCCCTGCTTCTGCTCTGACGGGGCCC GluGlyTrpPheAsnArgSerProTrpMetThrThrLeuLeuSerAlaLeuThrGlyPro	7440
7441	CTAGTAGTCCTGCTCCTGTTACTTACAGTTGGGCCTGCTTAATTAAAGGTTGGCC LeuValValLeuLeuLeuLeuThrValGlyProCysLeuIleAsnArgPheValAla	7500
7501	TTTGTAGAGAACGAGTGAGTCAGTCAGATCATGGTACTTAGGCAACAGTACCAAGGC PheValArgGluArgValSerAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGly	7560
7561	CTTCTGAGCCAAGGAGAAACTGACCTCTAGCCTCCAGTTCTAAGATTAGAACTATTAA LeuLeuSerGlnGlyGluThrAspLeuEnd	7620
7621	CAAGACAAGAACGAGGGAAATGAAAGGATGAAATGCAACCTAACCTCCAGAACCCAGG	7680
7681	AAGTTAATAAAAGCTCTAAATGCCCGAATTACAGACCCCTGCTGGCTGCCAGTAAATA	7740

Figure 3 cont.

19/22

7741 GGTAGAAAGGTACACACTTCCTATTGTTCCAGGGCCTGCTATCCTGGCCTAAGTAAGATAAC 7800
7801 AGGAAATGAGTTGACTAATCGCTTATCTGGATTCTGTAAAAGTGAATGGCACCATAGAAG 7860
7861 AATTGATTACACATTGACAGCCCTAGTGACCTATCTCAACTGCAATCTGTCACTCTGCC 7920
7921 AGGAGCCCACGCAGATGCGGACCTCCGGAGCTATTTAAAATGATTGGTCCACGGAGCGC 7980
7981 GGGCTCTCGATATTTAAAATGATTGGTCCATGGAGCGCGGGCTCTCGATATTTAAAAT 8040
8041 GATTGGTTTGTGACGGCACAGGCTTGTGTGAACCCCATAAAAGCTGTCCCGATTCCGCA 8100
8101 CTCGGGGCCGCAGTCCTCTACCCCTGGTGGTGTACGACTGTGGGCCCCAGCGCGCTTGG 8160
8161 AATAAAAATCCTTTGCTGTTGCATCAAAAAAAAAAAAAAAA 8209

Figure 4.

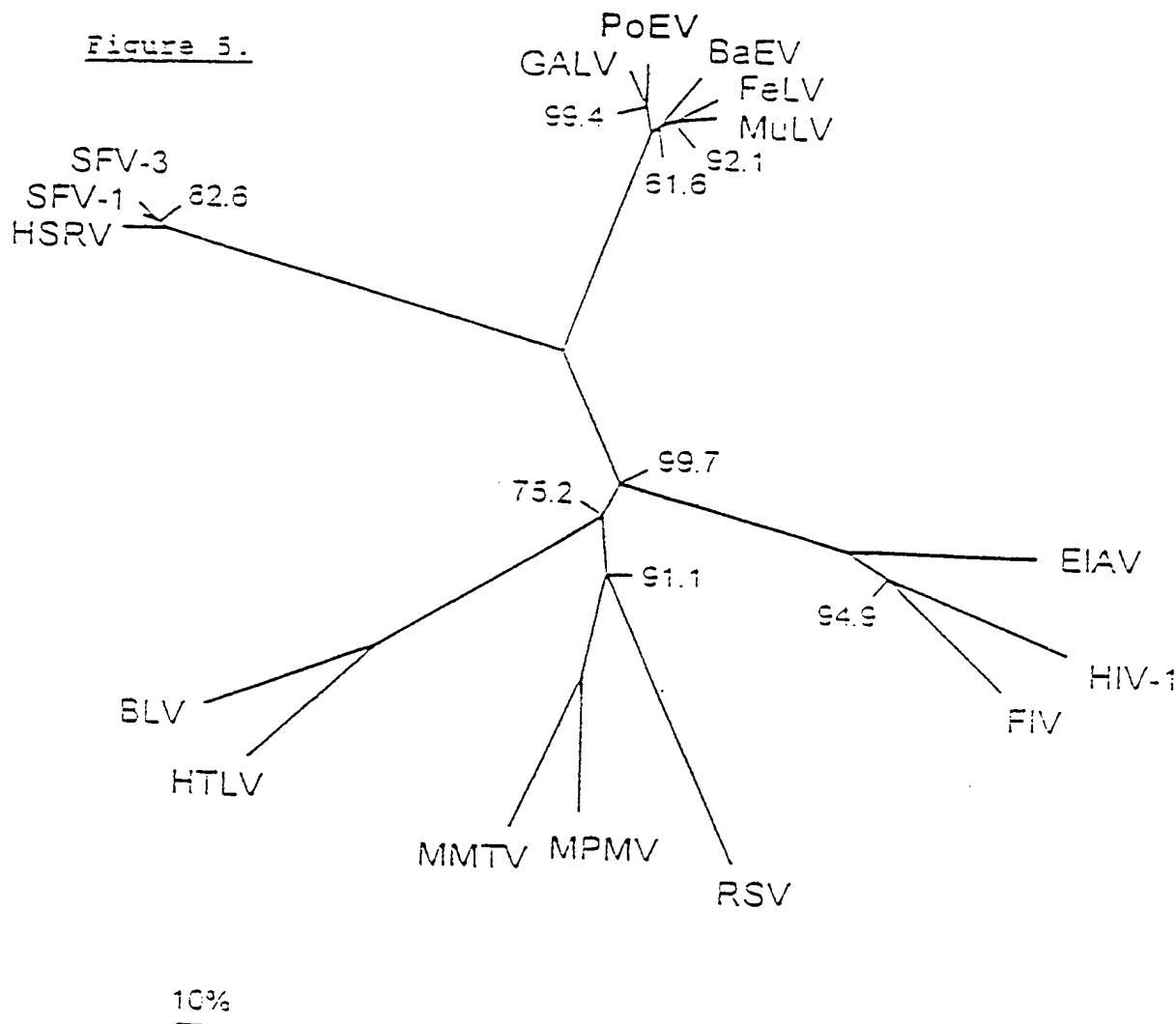
The same nucleotide sequence as represented by bases 5260 to 8210 in Figure 3 is also representative for this Figure, with the following changes:

<u>Position</u>	<u>Change</u>
5273	G-T
5341	C-T
5351	C-T
5353	T-C
5356	C-T
5426	G-A
5464	Insertion AGA
5607	C-T
5638	C-T
5792	T-C
6191	Insertion AA
6253	T-A
6255	Insertion A
6900	C-G

Such nucleotide changes result in the following amino acid changes in the ENV polypeptide.

<u>Position</u>	<u>Change</u>
7	R-W
192	R-K
193	Deletion
194	Deletion
197	Y-Q
198	S-E
199	K-N
200	V-I
201	D-Q
204	Y-I
205	E-N
206	Insertions: G, M, S
206	L-W
208	N-I
209	S-V
211	L-Y
212	L-K
427	F-L

21/22

Figure 5.

MuLV murine leukaemia virus
 FeLV feline leukaemia virus
 GaLV gibbon ape leukaemia virus
 SVV-1 simian foamy virus 1
 SFV-3 simian foamy virus 3
 HSRV human foamy virus
 SLV Bovine leukaemia virus
 HTLV human T-cell leukaemia virus
 MMTV murine mammary tumour virus
 MPMV Mason Pfizer monkey virus
 RSV Rous sarcoma virus
 FIV feline immunodeficiency virus
 HIV human immunodeficiency virus
 EIAV equine infectious anaemia virus

22/22

Figure 6

PPT | U3 OCT-1 c-Myb LyF-1 E47

1 AAGAAGTGGGGAATGAAAGGATGAAAATGCAACCTAACCCCTCCCAGAAC

ETS-1 AP-4

51 CAGGAAGTTAATAAAAAGCTCTAAATGCCCCCGAATTMCAGACCCTGCTG

NF-1 AP-1/TR

101 GCTGCCAGTAAATAGGTAGAAGGTCACACTTCCCTATTGTTCCAGGGCCTG

ETS-1/GATA GATA ETS-1 c-Myb AP-1 GATA

151 CTATCCTGGCCTAAGTAAGATAACAGGAAATGAGTTGACTAATCGCTTTAT

E47 AP-1

201 CTGGATTCTGTAAAACTGACTGGCACCATAGAAGAATTGATTACACATTG

AP-1 AP-1/GATA c-Myb AP-1

251 ACAGCCCTAGTGACCTATCTCAACTGCAATCTGTCACTCTGCCAGGAGC

E47 ETS-1 → CCAAT

301 CCACGCCAGATGCGGACCTCCGGAGCTATTTAAAATGATTGGTCCACCGGA

GATA → CCAAT ←

351 GCGCGGGCTCTCGATATTTAAAATGATTGGTCCATGGAGCGCGGGCTCT

GATA CCAAT ← AP-1/CREB

401 CGATATTTAAAATGATTGGTTGTGACGCACAGGGCTTGTGAACCC

TATA U3 | R

451 CATAAAAGCTGTCCCGATTCCGACTCGGGGCCGCAGTCCTCTACCCCTG

PADS polyA

501 CGTGGTGTACGACTGTGGGCCCCAGCGCGCTTGGAAATTTACCTCTTG

R | U5

551 CTGTTGCATCAAGACCGCTTCTYGTGAGTGATTGGGTGTCCGCCTCTT

U5 | PBS

601 CCGAKCCCGGACGAGGGGATTGTTCTTACTGGCCTTCATTGGTGC

651 GTTGGCCGGGAAATCCTGCGACC